Technical notes

How the identification of wine microorganisms has evolved within the AWRI Wine Microorganism Culture Collection

The Australian wine industry is fortunate to have access to a comprehensive repository of wine-related microorganisms (yeast, bacteria and fungi) that dates back more than 80 years. This collection both preserves Australia's winemaking heritage and supports the latest research on wine microbiology and molecular biology. Over time, as the collection has developed, the way microbial strains are identified has changed, adapting to the latest scientific knowledge and technologies. This article describes the evolution of techniques used for identification of microorganisms and how they have been adopted by the AWRI Wine Microorganism Culture Collection, as well as future directions.

Receival and identification of microbes

The AWMCC receives strains throughout the year, from both researchers and winemakers. When samples are received, the microbes of interest are isolated on selective agar and purified. Currently, yeast are identified using microscopy/morphology, physiology, biochemistry and DNA techniques, with results compared to databases and the closest match used as the species identity. Bacteria are identified via microscopy/morphology, physiology and biochemical testing. Limited molecular identification of certain bacteria types is also employed where a segment of the bacterial ribosomal DNA is sequenced.

Microorganism species identification is a dynamic area, and names are constantly being updated, so the collection is periodically updated with new species designations. The goal is to identify the isolates correctly and efficiently to make

What is the AWRI Wine Microorganism Culture Collection?

- The AWRI Wine Microorganism Culture Collection (AWMCC) is the largest repository of fermentation-associated yeast, fungi and bacteria in the southern hemisphere. It contains more than 5,000 different strains of microorganisms, mostly isolated from wine, with some Australian isolates dating back to the 1930s.
- The AWMCC provides the Australian wine industry with novel, non-commercially available yeast and bacterial strains for efficient and reliable fermentations and to shape and diversify wine style.
- The AWMCC is also essential for capturing the value of Australia's investment in microbial strain isolation, development, and characterisation. It contributes to the success of wine-related biological research projects, which depend on ready access to secure and correctly identified strains.

the process of screening yeast or bacteria for research projects more relevant, or to be able to provide appropriate cultures for winemakers.

Limitations of older identification methods

Historically, techniques such as microscopic analysis of cell size and shape (morphology), biochemical testing (e.g. Gram stain, catalase tests) and physiological testing (e.g. ability to ferment different carbon sources or ethanol tolerance) have been used to identify microorganisms. However, such morphological and biochemical techniques are increasingly being demonstrated to be prone to inaccuracies and biases in their interpretation, even when conducted by experienced investigators. Such techniques rely on comparing results to a reference, whether that be a large international database, or simply a type strain of an organism (usually the original isolate of that species) under the same conditions.

Use of DNA technologies

The development of DNA-based molecular techniques has allowed for many of the biases of traditional identification techniques to be eliminated. DNA-based technologies rely on either using the number and size of the chromosomes or the DNA sequence of the yeast or bacteria to identify the organism. DNA is the genetic code used by all living cells to translate information into proteins, which then determine the organism's observable traits (phenotype) including its physiology and morphology. It is much more efficient to analyse the DNA directly for identification rather than the outcome of several downstream steps that the DNA regulates.

In identifying yeast species, there have been several variations on DNA sequence-based techniques that have been used over many decades. Many have been used to determine the species present in a fermentation and some have been pioneered for wine microbe analysis at the AWRI. Early molecular techniques such as RFLP (restriction fragment length polymorphism), DNA/DNA association, CHEF, DGGE and AFLP gave different levels of identification of a species. They used different forms of (sometimes expensive) equipment, were time-consuming and had different sources of errors. Much like the classical tests, they remained at best a guide to the species present.

The advent of polymerase chain reaction (PCR) in the 1980s saw the development of a large number of techniques for the identification of DNA from cells. These techniques mostly focused on analysing amplified DNA (from various regions of the genome) that was then cut into fragments to perform the identification. These fragments were separated and visualised on agarose gels to create a pattern or fingerprint that is unique to each species (Figure 1).

While in most situations, these DNA fingerprinting techniques provide definitive species identification, they can be very dependent on technical skill, specific laboratory equipment and the conditions employed to create them. In addition, if there is no match of the sample fingerprint to an existing reference, then no accurate determination of species can be made.

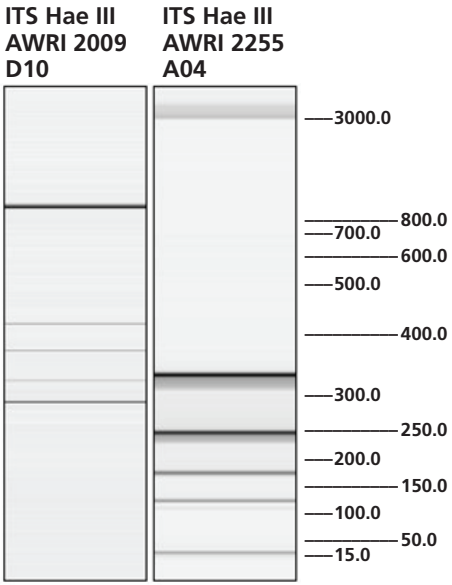

Figure 1. Restriction Fragment Length Polymorphism (RFLP) analysis of *Lachancea thermotolerans* (AWRI 2009) vs *Saccharomyces cerevisiae* (AWRI 2255)

Advantages of DNA sequencing

Unlike the fingerprinting techniques, DNA sequence analysis relies on determining the exact sequence of DNA bases (the units than make up DNA) that are present. DNA sequencing can be performed on both yeast and bacteria; however, the wine bacterium *Oenococcus oeni* is a recalcitrant organism for which analysis by sequencing can be challenging. Analysis of this bacteria at the AWRI is still performed mainly by biochemical and morphological analysis, but the number of bacteria that show these traits in a fermentation environment (wine, cider) are limited, so a positive identification is easily achieved. Nonetheless the ability to sequence the DNA of *Oenococcus oeni* strains in the future is an aim of the AWMCC and will eliminate the use of microscopy to identify the strains.

DNA sequencing can eliminate ambiguous results and confirm a correct identification; although, as for the fingerprinting approaches, the analysis and identification are only as good as the reference data available. Fortunately, databases of sequencing results from key

diagnostic regions in the DNA (for example ITS or D1/D2) have been built up over several years and thus a confident identification can generally be performed.

ITS sequencing analysis is dependent on PCR amplification of the internal transcribed spacer region (ITS) of the rDNA of yeast and fungi. This region is variable enough to provide a species-specific match, but not so variable as to prevent a species identification. Each yeast has a variation in ITS sequence size (the number of nucleotides in the region) which can be used as a quick verification of genus, but at times it is not enough to confirm the species. The sequencing of the ITS region is a standard method for identifying yeast species but other regions and genes are also used if the ITS database cannot differentiate the species. Sometimes there is a need to perform whole genome sequencing (WGS) to differentiate species for closely related species or complex genomes, such as the sequencing of the shipwreck yeast isolates (Figure 2) or the recent determination of three new species of *Aureobasidium* at the AWRI (Onetto et al. 2020).

The identification of wine isolates is a service that is performed when strains are deposited into the AWMCC. Collection staff have recently implemented a change in focus from RFLP analysis to direct sequencing of the yeast ITS region of the rRNA genes to allow unambiguous identification of yeast. This allows a direct online comparison to many more reference genomes without need of comparison to already confirmed species *in vitro*. The databases for ITS regions are vast and can identify most species found in wine environments.

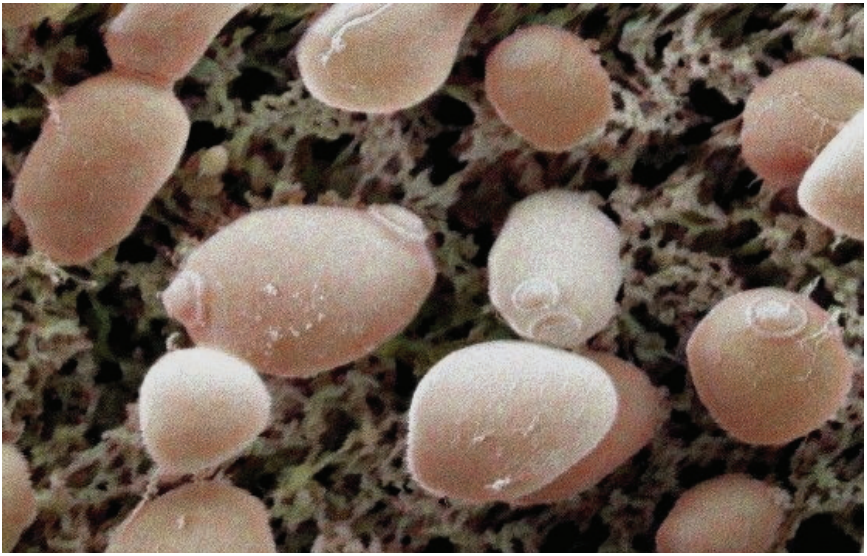

Figure 2. AWRI 3417 *Saccharomyces* shipwreck yeast (15,000x magnification)

Use of next generation sequencing technologies

Next generation sequencing (NGS), massively parallel or deep sequencing are related terms that describe a DNA sequencing technology that has revolutionised genomic research (Behjati et al. 2013). The advent of NGS technology in the sequencing of large numbers of species at one time means these technologies can be applied directly to samples of wines and ferments, where previously individual strains would have to be isolated on agar plates before the species could be identified (Borneman et al. 2013). These technologies are referred to as culture independent as they process the DNA present in the sample rather than extracts of DNA from pure cultures.

With yeast, for example, sample identities are categorised by attaching a known sequence of barcode DNA to the ITS sequences in the sample, thus allowing separation of each of the segments of ITS DNA based on the origin of the sample. For example, a winemaker could take 10 wine samples from each of five different vineyards and each sample would be given a specific DNA barcode ($10 \times 5 = 50$ samples) so there are 50 different barcodes, one for each sample. This is termed the DNA library. After processing, all the libraries can be pooled together for sequencing in one run, saving equipment running costs, technician time and consumables. Each wine sample or vineyard can then be filtered into groups (based on barcode information) depending on what information is needed (e.g. all yeast from one wine variety in five different vineyards).

Bioinformatics analyses are used to piece together these fragments by mapping the individual reads to the individual barcodes and thus relating all sequences back to the original sample. The bioinformatics pipelines (computational analysis programs) also refer all the sequence data to stored sequences in databases to achieve species identification. These methods are extremely sensitive and very high-throughput, which means that microbial communities can be more comprehensively characterised and many more samples can be tested than was possible with earlier methods (Sternes et al. 2017, Bokulich et al. 2012) (Figure 3).

The cost of sequencing DNA has come down significantly in the last 20 years, largely in response to the development of NGS and the bioinformatics used to analyse the results (Wetterstrand 2020). Prior to the introduction of NGS, the cost of DNA sequencing was reducing at a rate approximating 'Moore's law' (halving in cost every two years). Since its introduction around 2005, NGS has driven the per-base cost of DNA sequencing down even more quickly, from $US 5,292 per megabase in 2001 to less than US 1 cent per megabase in 2020 (Wetterstrand 2020).

The application of NGS in the analysis of fermentation has also allowed much greater understanding of the biodiversity of wine (Bokulich et al. 2012). The AWRI has used these technologies over the last few years to greatly increase the depth of knowledge of certain fermentations, which would not have been possible at the cost or sensitivity previous technology allowed. Examples include an investigation of the yeast and bacteria present in wild-fermented cider gum sap in Tasmania (Varela et al. 2020) and studies on the impact of sulfur dioxide on the yeast species present uninoculated grape juice fermentations (Cuijvers et al. 2020).

The possibility of using NGS for the identification of many species of yeast or bacteria in the AWMCC is currently being investigated. Unlike the analysis of a wine samples, however, each strain must be re-cultured from frozen stocks, checked for purity and have DNA extracted and barcoded. This is traditionally a labour-intensive process, but may allow a single sequencing run to identify the microorganisms (as yeast and bacterial DNA can be sequenced in the same run of pooled DNA).

To deal with the high labour investment for NGS, the AWRI has purchased a robotic liquid handling robot (epMotion 5075) which combines automated liquid handling and the software needed for NGS library preparation. This machine can take much of the labour and monotonous work out of NGS sample preparation which in turn helps minimise errors.

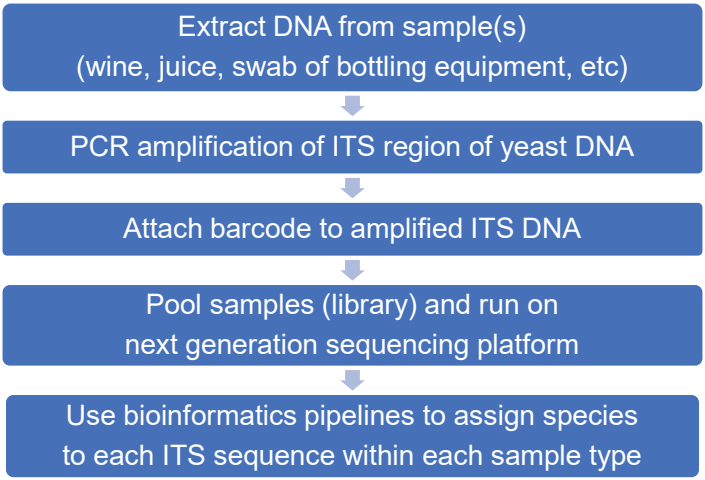

Figure 3. Next generation sequencing flow chart for yeast identification in wine samples

Conclusion

Identification of yeast and bacteria in the AWMCC has evolved over many decades since its inception, from traditional biochemical and morphological analysis of pure cultures, to pattern-based molecular fingerprinting methods in the last three decades, and finally to DNA sequencing more recently. This progression of identification techniques has reduced the ambiguity of the identification and the time needed to identify a species with reliability. With the advent of robotics and higher quality next generation sequencing, the ability to identify hundreds of yeast and bacteria from a single sample is now possible in a cost-effective manner.

For more information about the strains available or depositing microorganisms in the collection, please visit the AWMCC website: www.awri.com.au/research_and_development/wine-microorganism/winemaking-yeast-and-bacterial-strains/.

Simon Dillon – Scientist, *simon.dillon@awri.com.au*

References

Onetto, C.A., Schmidt, S.A., Roach, M.J., Borneman, A.R. 2020. Comparative genome analysis proposes three new *Aureobasidium* species isolated from grape juice. FEMS Yeast Res. 20(6): foaa052.

Borneman, A., Herderich, M., Johnson, D. 2013. The DNA of innovation. Wine Vitic. J. 28 (5): 52–56.

Behjati, S., Tarpey, P.S. 2013. What is next generation sequencing? Arch. Dis. Child. Educ. Pract. 98: 236–238.

Bokulich, N.A., Joseph, C.M.L., Allen, G., Benson, A.K., Mills, D.A. 2012. Next-generation sequencing reveals significant bacterial diversity of botrytized wine. PLoS ONE 7(5): e36357.

Cuijvers, K., Van Den Heuvel, S., Varela, C., Rullo, M., Solomon, M., Schmidt, S., Borneman, A. 2020. Alterations in yeast species composition of uninoculated wine ferments by the addition of sulphur dioxide. Fermentation 6(2): 62.

Sternes, P.R., Lee, D., Kutyna, D.R., Borneman, A.R. 2017. A combined meta-barcoding and shotgun metagenomics analysis of spontaneous wine fermentation. GigaScience 6 (7): 1–10.

Varela, C., Sundstrom, J., Cuijvers, K., Jiranek, V., Borneman, A. 2020. Discovering the indigenous microbial communities associated with the natural fermentation of sap from the cider gum *Eucalytpus gunnii*. Sci. Rep. 10: 14716.

Wetterstrand, K.A. 2020. DNA sequencing costs: data from the NHGRI genome sequencing program (GSP) Available from: www.genome.gov/sequencingcosts. Accessed 4 March 2021.